Chains of chirplets for the detection of gravitational wave chirps

Éric Chassande-Mottin[1], Archana Pai [2] and Olivier Rabaste [1]
[1] CNRS, AstroParticule et Cosmologie, 10, rue Alice Domon et Léonie Duquet,
75205 PARIS Cedex 13, FRANCE
[2] Max-Planck Institut für Gravitationsphysik, Am Mühlenberg 1,
14476 Potsdam, GERMANY

ABSTRACT

A worldwide collaboration attempts to confirm the existence of *gravitational waves* predicted by Einstein's theory of General Relativity, through direct observation with a network of large-scale laser interferometric antennas. This paper is a contribution to the methodologies used to scrutinize the data in order to reveal the tiny signature of a gravitational wave from rare cataclysmic events of astrophysical origin. More specifically, we are interested in the detection of short frequency modulated transients or gravitational wave chirps. The amount of information about the frequency vs. time evolution is limited: we only know that it is smooth. The detection problem is thus non-parametric. We introduce a finite family of "*template waveforms*" which accurately samples the set of admissible chirps. The templates are constructed as a puzzle, by assembling elementary bricks (the chirplets) taken a dictionary. The detection amounts to testing the correlation between the data and the template family. With an adequate time-frequency mapping, we establish a connection between this correlation measurement and combinatorial optimization problems of graph theory, from which we obtain efficient algorithms to perform the calculation. We present two variants. A first one addresses the case of amplitude modulated chirps and the second allows the joint analysis of the data from several antennas. Those methods are not limited to the specific context for which they have been developed. We pay a particular attention to the aspects that can be source of inspiration for other applications.

Keywords: Gravitational waves, detection, chirp signals, chirplets, dynamic programming

1. CONTEXT AND MOTIVATION

1.1 Gravitational wave astronomy: the dawn of a new era of astrophysics

The Theory of General Relativity predicts the existence of wave solutions to the equations which govern the dynamics of space-time. These solutions are referred to as *gravitational waves*[1] (GW) and can be regarded as propagating deformations of space-time. An international collaboration currently attempts to confirm this prediction by the direct observation with a network of GW antennas (see Fig. 1).

Although these antennas have different sensitivities due to different scales and technologies, they all follow the same measurement principle. Since GW acts locally as a strain i.e., it causes a fractional change of the distance between two points with constant coordinate separation*, the idea is to use high precision laser metrology to measure this effect. GW antennas are Michelson-Morley interferometers.[2] In a nutshell, coherent twin laser beams are propagated back and forth along two perpendicular directions i.e., the two arms (of length L) of the interferometer. The two beams eventually combine and interfere. From the measurement of the resulting light intensity, we obtain the difference δL of their optical path lengths and thus the fractional length change $\delta L/L$ due to an impinging GW.

The observation of GW requires an antenna and also a generator. GW are produced by accelerated masses, the same way electromagnetic waves are produced by accelerated charges. However, space-time is very hard to

Send correspondence to Éric Chassande-Mottin
Éric Chassande-Mottin: E-mail: ecm@apc.univ-paris7.fr
*Those two points can be physically enacted by freely falling particles.

deform like a stiff membrane. Very large masses and relativistic velocities are required to generate observable GW. Only astrophysical systems possess those characteristics and even for these systems, the GW strain amplitude is expected to be tiny, of the order of $\delta L/L \sim 10^{-21}$.

GW antennas are designed to reach this target sensitivity. First, they have long arms to relax as much as possible the requirement on the precision for the measurement of the differential length δL. With L in the realistic kilometric scale, the specification becomes $\delta L \sim 10^{-18} m$. As a comparison, this is of the order of one-thousandth the "diameter" of the proton. To achieve this unprecedent accuracy, GW antennas use a more sophisticated measurement scheme than the one explained above and they include innovative technologies developed specifically. To highlight few of those, the instrument is entirely isolated from the ground by advanced seismic attenuators. The whole instrument operates under ultra high vacuum. For this reason, it is embedded in a very large vacuum chamber (these chambers are among the world's largest). GW antennas are equipped with high quality optics among the best mirrors in terms of losses and reflectivity. This allows an observation window in the bandwidth ranging from few 10 to few 10^3 Hz (see Fig. 2).

Several source candidates have been identified in this frequency window (see Ref. 3 for a complete overview). Two types can be distinguished: long-term/weak amplitude (which includes periodic waves produced by rapidly spinning neutron stars, or the stochastic background created in the early Universe) and short-term/large amplitude (associated to cataclysmic phenomena, such as supernova explosions or the coalescences of neutron star or black hole binaries). Here, we are interested in the second category, which consists in GW transients lasting typically few hundreds of milliseconds to few seconds. The next section details and motivates the specific model of GW transients we consider.

The farther the source, the weaker the GW signal. The sensitivity of the GW antennas limits their observational field to a finite horizon (with typical distance of few tenths of Mpc). Within this volume of Universe, the projections of the source populations predict small numbers of individuals. Consequently, the search consists in detecting rare transients (with a typical rate of occurrence of a few per year) in the time series acquired during long period of observation (lasting typically a year).

The scope of GW observation goes beyond the only verification of their existence. Observing GW give access to information about the physics of their sources that are difficult to access otherwise (with an optical telescope for instance). The first detection will inaugurate the new era of *gravitational wave astronomy*.

1.2 GW chirps as generic signatures of orbiting or spinning sources

The dynamics of the source drives the GW emission. When the source is a slowly moving[†] orbiting or spinning system in a plane (making an angle ϵ with the line of sight, see Fig. 3), the rotational motion inducts a quasi-periodic and frequency modulated GW emission. A GW is characterized by two independent polarizations h_+ and $h_\times$ which we write using the following compact form, viz.

$$h(t) \equiv h_+(t) + i h_\times(t) = \frac{1+\cos^2\epsilon}{2}\ A(t)\cos\varphi(t) + i\cos\epsilon\, A(t)\sin\varphi(t), \tag{1}$$

where the amplitude $A(t) > 0$ and phase $\varphi(t)$ are related to the orbital or spinning phase. We refer to this signal as *GW chirps*. Note that, the two polarizations are coupled in this model and this results from the assumptions describing the source.

Several sources of GW chirps are well-known. One example is the coalescence of compact object (neutron stars or black holes) binaries. The models of the coalescence predict the time evolutions of chirp phase and amplitude (power laws in a first approximation) with great accuracy. Currently, the efforts are mainly concentrated on the detection of GW chirps with these specific phase and amplitude.

Our objective is different. The model in Eq. (1) is a robust prediction and remains valid whenever the motion is rotational/orbital independent of the physical phenomena which determines it. In other words, GW chirps are generic signatures of orbiting/spinning systems. This is the basic motivation for searching for wider ranges of GW chirps.

[†]The velocities are small with respect to the speed of light.

In physically realistic situations, the phase evolution is however not totally arbitrary: the phase $\varphi(t)$, the frequency $f(t) \equiv (2\pi)^{-1} d\varphi/dt$ and its derivatives are expected to be continuous. Also, the variations of the frequency are directly related to the amount of available energy. The energy being finite, those variations are bounded. In this work, we only consider the bounds on the modulation rates $|df/dt|$ and $|d^2f/dt^2|$. Similarly, we expect the amplitude $A(t)$ to be a smooth function with a limited variation rate. These arguments set constraints on the regularity of the amplitude and phase evolutions.

Besides those regularity constraints, we do not have other information about the GW chirp. As opposed to the detection of coalescing binary chirps, we do not have an analytical expression describing the phase and amplitude at our disposal. The detection of generic GW chirp is thus a non-parametric problem. Several methods have been proposed to address this problem which include the signal track search[4] (STS) , the best chirplet path search[5] and the best chirplet chain (BCC) search[6] proposed by the authors. In this article, we present the basic ideas of the BCC method. In the next section, we start with the simple case where the phase only is unknown (the amplitude is assumed to be constant). In Sec. 3, we discuss extensions and variants[7,8] of this first case.

2. NON-PARAMETRIC DETECTION OF GW CHIRPS

2.1 Basic ideas

The GW antenna receives a linear combination of the two polarizations. Concretely, we have $\delta L/L = F_+ h_+ + F_\times h_\times$ where the antenna pattern function F_+ and $F_\times$ depends on the relative orientation of the incident GW and the detector plane.

Therefore, let $s(t) \equiv F_+ h_+(t) + F_\times h_\times(t)$ be the signature left in the detector output by an incoming GW chirp. Our objective is to detect this signature in the (additive) instrumental noise.

Figure 2 displays the noise spectrum of the main GW antenna. The noise spectrum is not flat but follows a curve whose global shape is common to all detectors (steep increase at the lower and higher ends of the observation window). We assume that the raw detector output is pre-conditioned by a whitening filter. We denote the resulting time-series $x(t)$ and consider that $s(t)$ designates the "whitened" version of the actual GW signal.

For GW chirp as defined in Eq. (1), the signature can be conveniently written as $s(t) \equiv \Re\{\bar{F}^* A(t) \exp i\varphi(t)\}$ where $\bar{F} = (1 + \cos^2\epsilon)/2\ F_+ + i \cos\epsilon\ F_\times$ is the complex antenna pattern. Here, z^* designates the complex conjugate of z. This expression evidences that the signature $s(t)$ is a scaled and phase shifted version of the complex signal $A(t) \exp i\varphi(t)$ which depends only on the "dynamical" part of the GW.

If the chirp phase and amplitude were known in advance, the detection would optimally be performed with the likelihood ratio. Under Gaussian assumption for the noise, this likelihood ratio essentially reduces to the *quadrature matched filtering* statistic,[9]

$$\ell_p(x) = \left| \sum_{k=0}^{N-1} x(t_k) s_p^*(t_k) \right|^2 , \tag{2}$$

where we work with blocks of N data samples taken at discrete times $t_k = k/f_s$ with a sampling frequency $f_s \sim$ few kHz to cover the detector band. The duration $T = N/f_s$ corresponds to the lifetime of a GW chirp in the detector band, i.e., typically $T \sim$ few 10^{-1} to 1 sec, so that $N \sim$ few 10^3 to 10^4.

The above statistic amounts to correlating the data with the reference waveform $s_p(t)$ which is a copy of $s(t)$ normalized to unit norm. More precisely, we have $s_p(t) = a(t) \exp i\varphi(t)$ where the amplitude $a(t) \propto A(t)$ is rescaled to unit norm i.e., $\sum_{k=0}^{N-1} a(t_k)^2 = 1$. The subscript $p \equiv \{A(\cdot), \varphi(\cdot)\}$ stresses the dependency on the amplitude and phase functions taken here as parameters. The result of the correlation is then compared to a threshold to decide whether the chirp is present or not in the data.

In reality, the chirp phase and amplitude are not known. Inspired by the principles of the *generalized likelihood ratio test*[9] (GLRT), we apply the statistic to all possible waveforms and check if at least one shows a significant correlation with the data. The statistic is thus

$$\mathcal{L}(x) = \max_{p \in \mathcal{P}} \ell_p(x), \tag{3}$$

where $\mathcal{P}$ denotes the set of all admissible phases and amplitudes i.e., which satisfy the regularity conditions defined in Sec. 1.2.

The above maximization problem is non-linear and non-convex and it is performed over a continuous set of waveforms. This makes the problem difficult to solve in this form. These difficulties can however be circumvented by discretizing $\mathcal{P}$ with a finite family $\tilde{\mathcal{P}}$ of *template waveforms*. The maximization over the continuous set $\mathcal{P}$ is replaced by the search over $\tilde{\mathcal{P}}$ which can be performed numerically.

The templates are constructed as a "puzzle", by assembling elementary bricks taken from a dictionary. The building blocks are small pieces of chirp signal, or *chirplets* which we assemble in connected *chains of chirplets*. In Sec. 2.2, we define those objects in more details. This construction scheme allows to generate a large variety of chirps although the number of building blocks is limited. The dictionary and the assembling rules are chosen such that any element of the admissible chirp set $\mathcal{P}$ can be closely approximated.

In Ref. 5, the authors rather speak of *chirplet graph* which turns out to be a better terminology since it makes a connection with graph theory and related combinatorial optimization problems. Indeed, the representation of the template as a chain or a graph allows the decomposition of the whole and difficult maximization problem in Eq. (3) into an interweaving of smaller and simpler problems. This gives rise to an efficient scheme which yields the global maximum recursively from the sequence of the smaller problem solutions. We detail this point in Sec. 2.3.

2.2 Chains of chirplets: templates for GW chirps

We now elaborate on the ideas given in the previous section restricting to the case where only the phase $\varphi(t)$ is unknown and the amplitude $A(t) = A$ is assumed to be constant. We thus have $p = \{\varphi(\cdot)\}$ and $s_p(t) = N^{-1/2} \exp i\varphi(t)$ (i.e., $a(t) = \text{Cst} = N^{-1/2}$ due to the unit norm rescaling).

The construction of the chirplet dictionary is easily represented in the time-frequency (TF) plane as displayed in Figure 4. We define a TF grid by dividing the time and frequency axes into N_t and N_f intervals respectively. We call *chirplet* a TF segment joining two successive grid vertices at whose frequency doesn't vary more than $\pm N_r'$ frequency bins over a time bin. We form a *chirplet chain* (CC) and impose its continuity. We also impose that the slopes of two successive chirplets in the chain do not vary more than $\pm N_r''$ frequency bins. The values of the regularity parameters N_r' and N_r'' are fixed by the bounds on the frequency modulation rates.

In a sense, the family $\tilde{\mathcal{P}}$ of all CC "samples" the set $\mathcal{P}$ of all admissible chirps. An evaluation of the accuracy of this sampling can be obtained. We define the "distance" between an arbitrary chirp of phase $p = \{\varphi(\cdot)\}$ and the CC of phase $\tilde{p} = \{\tilde{\varphi}(t)\}$ by $\delta(p, \tilde{p}) \equiv \ell_p(s_p) - \ell_{\tilde{p}}(s_p)$. We get the distance between $\mathcal{P}$ and $\tilde{\mathcal{P}}$ from the following worst case estimate[6]

$$\sup_{p \in \mathcal{P}} \inf_{\tilde{p} \in \tilde{\mathcal{P}}} \left(\frac{|\delta(p, \tilde{p})|}{\ell_p(s_p)} \right)^{1/2} \sim \mathcal{O}\left(\frac{1}{N_t^2} + \frac{1}{N_f} \right). \tag{4}$$

Therefore, for N_t and N_f large enough, we can safely replace $\mathcal{P}$ by $\tilde{\mathcal{P}}$ in Eq. (3) which allows to perform the maximization numerically. In the next section, we examine which search algorithm can be used.

Note that we choose a regular TF grid for simplicity. There are however theoretical motivations for changing to a multi-scale dictionary as developed in Ref. 5. We won't investigate this possibility here.

2.3 Searching through the chirplet chain family

We want to maximize $\ell_{\tilde{p}}(x)$ over the finite family $\tilde{\mathcal{P}}$. The cardinal of $\tilde{\mathcal{P}}$ grows rapidly (exponentially) with N_t (and thus with N). For realistic signal duration, it becomes so large that the use of an exhaustive search through $\tilde{\mathcal{P}}$ to find the maximum is computationally prohibitive. However, the structure of chirplet chains offers the possibility to use efficient alternatives to the exhaustive search borrowed from combinatorial optimization.

Combinatorial optimization problems are much simpler to solve when the objective function is *additive* i.e., when it can be written as the sum of contributions depending on one variable only. In the case of interest here, the overall square modulus in the expression of $\ell_{\tilde{p}}$ (see Eq. (2)) prevents from expressing this function as a sum of terms involving one chirplet only. Therefore, the objective function is *not* additive strictly speaking, but as we will show now, it is *close* to an additive function.

We first reformulate $\ell_{\tilde{p}}$ in the TF plane using the discrete Wigner-Ville (WV) distribution. The WV distribution is defined as[10]

$$w_x(t_n, f_m) \equiv \sum_{k=-k_n}^{+k_n} x(t_s\lfloor n + k/2\rfloor)x^*(t_s\lfloor n - k/2\rfloor)\, e^{-2\pi imk/(2N)}, \tag{5}$$

where $k_n \equiv \min\{2n, 2N - 1 - 2n\}$ and $\lfloor\cdot\rfloor$ defines the integer part. The arguments of WV are the discretized time $t_n = n/f_s$ and frequency $f_m = f_s m/(2N)$ with $n = 0, \ldots, N-1$ and $m = 0, \ldots, 2N-1$.

The WV distribution combines two key properties which allow to reformulate the quadratic matched filter as time-frequency path integral.[11] First, it is *unitary*[10] which means that the scalar product of two time-series can be rewritten in terms of the scalar product of their TF distributions via the *Moyal formula*. Applying this property to Eq. (2), we can re-express the statistic as the correlation of the WV of the data with the template WV, viz.

$$\ell_{\tilde{p}}(x) = \frac{1}{2N}\sum_{n=0}^{N-1}\sum_{m=0}^{2N-1} w_x(t_n, f_m)w_{s_{\tilde{p}}}(t_n, f_m). \tag{6}$$

Second, the WV distribution is geometrically simple for chirps: it is almost Dirac along this TF curve, viz.

$$w_{s_{\tilde{p}}}(t_n, f_m) \approx 2\delta(f_m - \tilde{f}(t_n)), \tag{7}$$

where $\tilde{f}(t) = (2\pi)^{-1}d\tilde{\varphi}/dt$ is the instantaneous frequency of the CC. This approximation can be written as the sum of contributions due to individual chirplets, namely

$$w_{s_{\tilde{p}}}(t_n, f_m) \approx 2\sum_{j=0}^{N_t-1} \delta(f_m - \tilde{f}_j(t_n))\Pi_{\tilde{\mathcal{T}}_j}(t_n), \tag{8}$$

where $\tilde{f}_j(t)$ is the frequency of the j^{th} chirplet and $\tilde{\mathcal{T}}_j = \{t = n/f_s \text{ for } jb \le n < (j+1)b\}$ denotes its time support. Here, $b \equiv |\tilde{\mathcal{T}}_j| = N/N_t$ is the number of samples per chirplet. The function Π designates the boxcar function i.e., $\Pi_T(t) = 1$ when $t \in T$ and $\Pi_T(t) = 0$ otherwise.

Clearly, the approximation amounts to neglecting the so-called *interferences*.[10] The interferences are cross-terms in the WV which results from the quadratic interaction between the components of the signal. In the present case, the components coincide with the chirplets. Combining Eqs. (6) and (7), we obtain an approximation of the statistic $\hat{\ell}_{\tilde{p}}(x)$

$$\ell_{\tilde{p}}(x) \approx \hat{\ell}_{\tilde{p}}(x) \equiv \frac{1}{N}\sum_{n=0}^{N-1} w_x(t_n, \tilde{f}(t_n)) = \frac{1}{N}\sum_{j=0}^{N_t-1}\sum_{t_n\in\tilde{\mathcal{T}}_j} w_x(t_n, \tilde{f}_j(t_n)), \tag{9}$$

which reduces to the path integral of the data WV along the CC frequency.

The approximated statistic *is* an additive function. In other words, we eliminated the non-additive part of $\ell_p(x)$ i.e., the cross contributions between chirplets. Since $\hat{\ell}_{\tilde{p}}$ is additive, its maximization falls into the category of problems akin to the *shortest path problem*[12‡] that dynamic programming resolves in polynomial time. We refer to this TF pattern algorithm as *BCC search*.

Figure 5 illustrates the efficiency and robustness of the method. We apply the method to search a "random" chirp (i.e., a chirp randomly chosen in $\mathcal{P}$). We see that the frequency estimation fits quite nicely the injected signal. The interested reader may find in Ref. 6 a thorough evaluation of the performances of the algorithm.

3. VARIANTS AND EXTENSIONS

In this section, we discuss two extensions of the BCC search presented previously.

‡Note that we are rather searching here for the "longest" path and that the length is measured as a Lebesgue type integral over w_x.

3.1 Detecting GW chirps with a varying envelope

In Sec. 2.1, we proposed a method for the detection of GW chirps of constant amplitude. The method relies on two ingredients namely, (i) the discretization of the GW chirp set by a family of template waveforms obtained by chaining chirplets taken from a dictionary and (ii) an efficient time-frequency algorithm to search the template family. These two ingredients can be also obtained when the chirp amplitude is *not* constant (i.e., the GW chirp is modulated in amplitude), provided few adaptations.

The dictionary used when the GW chirp has a constant amplitude, contains chirplets of equal amplitude. In the next section, we change the dictionary by introducing the possibility to have chirplets with different amplitudes.

3.1.1 Chirplet chains with amplitude modulation

We now consider a dictionary of chirplets with constant but different amplitudes as illustrated in Fig. 4. Chains of chirplets taken in this dictionary have a piece-wise constant amplitude. We discretize the range of the chirplet amplitudes by a dyadic scale of N_a bins, namely the chirplet amplitudes are taken in $\{\tilde{a}_l \text{ for } l = 0, 1, \ldots, N_a - 1\}$ with $\tilde{a}_l = 2^{-l/2}/\sqrt{b}$ for $l < N_a - 1$ and $\tilde{a}_{N_a-1} = 0$. We control the regularity of the amplitude chain by imposing that the amplitude index l of two successive chirplets in a chain does not differ by more than the regularity parameter $\pm N_s'$.

The reference waveform $s_p(t) = a(t) \exp i\varphi(t)$ in Eq. (2) is normalized to unit norm i.e., $\sum_{k=0}^{N-1} a^2(t_k) = 1$. The chirplet chain $s_{\tilde{p}}(t) = \tilde{a}(t) \exp i\tilde{\varphi}(t)$ has to satisfy this normalization as well. This implies the following normalization constraint for the sequence of the chirplet amplitudes. Let l_j be the amplitude index of the j^{th} chirplet, we must have

$$b \sum_{j=0}^{N_t-1} \tilde{a}_{l_j}^2 = \sum_{\substack{j=0 \\ l_j \neq N_a-1}}^{N_t-1} 2^{-l_j} = 1. \tag{10}$$

We see here that the choice of the dyadic scale for the amplitudes ensures that, even if the number of possible amplitudes is limited, there is a large diversity of sequences which satisfy the above normalization constraint.

The duration or time support (i.e., number of the non-vanishing samples) of the chain and the peak of the amplitude sequence are linked by the normalization constraint. For instance, if a chain has one chirplet with $l = 0$ (i.e, $\tilde{a} = 1/\sqrt{b}$), the amplitudes of all the other chirplets must be zero (i.e., $\tilde{a} = 0$ and $l = N_a - 1$). The longest time support is determined by the value of the smallest amplitude. In other words, the choice of N_a fixes the largest number of non-vanishing chirplets that a chain can have. When $N_a = \log_2(N_t) + 2$, this number is maximum and equal to N_t such that the chains can last the whole time interval T with an amplitude $\tilde{a}_{N_a-2} = N^{-1/2}$.

3.1.2 Joint search over amplitude and frequency

Similarly to Sec. 2.3, the template family generated from the above dictionary is too large to be searched through exhaustively and thus requires the use of alternative algorithms. Here, we change the TF path search proposed earlier to fit the case where the chirplet amplitude evolves. In particular, this algorithm relies on the approximation of the template WV $w_{s_{\tilde{p}}}(t, f)$ given in Eq. (7) which has to be adapted to present case.

The compatibility of the WV with the multiplication allows us to write the template WV in terms of the WV of the amplitude and phase separately, viz.

$$w_{s_{\tilde{p}}}(t_n, f_m) = \frac{1}{2N} \sum_{m'=0}^{2N-1} w_{\tilde{a}}(t_n, f_{m'}) w_{\tilde{e}}(t_n, f_m - f_{m'}), \tag{11}$$

where $\tilde{e}(t) \equiv \exp i\tilde{\varphi}(t)$. Eq. (7) provides a model for $w_{\tilde{e}}(t, f)$. We approximate $w_{\tilde{a}}(t, f)$ as follows

$$w_{\tilde{a}}(t_n, f_m) \approx 2N\, \tilde{a}(t_n) \Pi_{\mathcal{F}(t_n)}(f_m), \tag{12}$$

where $\mathcal{F}(t_n)$ designates a frequency bandwidth which needs to be determined.

The idea here is to mimic the characteristic features of the WV of a low-pass function: the main contributions are concentrated about the axis $f = 0$ and the width $\mathcal{F}$ along the frequency axis is inversely proportional to the time support (due to the Heisenberg inequality). Owing to Eq. (10), large amplitude implies short time support and thus large frequency bandwidth. This argument suggests to choose $\mathcal{F}(t_n) = \{f \text{ such that } |f| \leq 2N\tilde{a}(t_n)/f_s\}$.

Combining Eqs. (11) and (12), and splitting the result in a sum of contributions from each chirplet, we get

$$w_{s_{\tilde{p}}}(t_n, f_m) \approx 2N \sum_{j=0}^{N_t-1} \tilde{a}_{l_j} \Pi_{\tilde{\mathcal{F}}_j}(f_m - \tilde{f}_j(t_n)) \Pi_{\tilde{\mathcal{T}}_j}(t_n), \tag{13}$$

where $\tilde{\mathcal{F}}_j = \{f \text{ such that } |f| \leq 2N\tilde{a}_{l_j}/f_s\}$ is the frequency width of the template WV associated to the j^{th} chirplet.

Replacing this expression in Eq. (6), we obtain the following approximation of the statistic

$$\hat{\ell}_{\tilde{p}}(x) = \sum_{j=0}^{N_t-1} \sum_{t_n \in \tilde{\mathcal{T}}_j} \sum_{f_m - \tilde{f}_j(t_n) \in \tilde{\mathcal{F}}_j} \tilde{a}_{l_j} w_x(t_n, f_m), \tag{14}$$

which has to be maximized jointly over the amplitude and phase/frequency. This maximization problem is similar to the one we end up with in Sec. 2.3. The difference is that we integrate along a two-dimensional time-frequency stripe of adjustable width and height instead of a one-dimensional path of fixed height.

Again, the objective function is additive but a major difference is that the maximization is subject to the normalization constraint in Eq. (10). This leads to the class of constrained optimization problems akin to the *change making problem* and other variations of the *knapsack problem*.[12] Again, a procedure based on dynamic programming solves this problem in polynomial time. Note that, due to the larger complexity of the problem, the algorithm is significantly more demanding in terms of number of numerical operations and memory space than the one proposed for the case where the amplitude is constant.

We present in Fig. 6 an illustration of the method with a "random" chirp. We see that, while the time support is somewhat underestimated, the central part of the chirp is well identified.

3.2 Detecting GW chirps using multiple antennas

The operational schedule of the global network of GW antenna is arranged in such a way that always, at least two antennas take data. Several improvements can be gained from this joint observation. First, the signal-to-noise is improved by combining the information received by each antenna. This results in an increase of the sight horizon and thus of the number of observable sources. Second, the wide aperture of GW antenna prevents the accurate identification of the sky position of the source. The observation with multiple antennas allows (when they are sufficiently distant) to locate the source with a much better accuracy. This provides an precious information for an eventual follow-up with other instruments like an optical telescope.

An impinging GW chirp produces the same signature (up to a scaling and phase factor) in the output of each antenna. Let $s^d(t) = \Re\{\bar{F}^{d*} A(t-\tau^d) \exp i\varphi(t-\tau^d)\}$ be the signature received by the d^{th} antenna of the network counting N_d detectors. The antenna pattern $\bar{F}^d$ depends on the alignment of the antenna and the source. The angles which define their relative orientation (see Fig. 3) are the spherical coordinates of the source (i.e., the co-latitude θ and longitude ϕ) and the polarization and inclination angles ψ and ϵ. (Note that these angles are *a priori* unknown parameters that have to be estimated from the data). In the above expression of the signature, the time t is measured by the clock of a (arbitrary) fiducial observer. The time delay τ^d measures the difference in the time of arrival at the antenna d with respect to the fide. It depends on the direction of arrival of the wave and hence on θ and ϕ.

The ideas and techniques of sensor array processing (for a recent review, see Ref. 13) apply in the present context. We consider the network of antennas as a sensor array which produces a multidimensional output $\mathbf{x}(t) = [x^1(t+\tau^1) \dots x^{N_d}(t+\tau^{N_d})]^T$. This output is constructed as a column vector which stacks the data from each detectors. A time-shift is applied to compensate the time delays assuming a certain source sky location.

With this formalism, the signature received by the network is $\mathbf{s}(t) = \Re\{\bar{\mathbf{F}} A(t) \exp i\varphi(t)\}$ where $\bar{\mathbf{F}} \equiv [\bar{F}^1 \dots \bar{F}^{N_d}]^T$. The GW signature thus belongs to the 2-dimensional vector space defined by $\bar{\mathbf{F}}$ and $\bar{\mathbf{F}}^*$ which we refer to as *GW polarization plane*.

If we consider the simplified case where the instrumental noise is Gaussian and white, the likelihood ratio associated to the detection of the GW signature can be written as the combined quadrature matched filtering of two *synthetic streams* $y_l(t)$ (with $l = 1$ and 2) which linearly combines the data from all the detectors,[7] viz.

$$\ell_p(\mathbf{x}) = \left| \sum_{k=0}^{N-1} y_1(t_k) s_p^*(t_k) \right|^2 + \left| \sum_{k=0}^{N-1} y_2(t_k) s_p^*(t_k) \right|^2 . \tag{15}$$

The synthetic streams are projections of the data onto the GW polarization plane, namely $y_l(t) = \mathbf{u}_l^{T*} \mathbf{x}(t)$ where the vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ form an orthonormal basis of this plane.

Synthetic streams are akin to beam-former solutions.[13] The vectors $\mathbf{u}_l$ acts as "spatial" filters which adjust the phase of the output of each antenna such that the signatures they carry add up constructively.

To maximize $\ell_p(\mathbf{x})$ over the admissible chirp set $\mathcal{P}$, we use the template families $\tilde{\mathcal{P}}$ defined in Sec. 2.2 or Sec. 3.1 depending on the assumption made on the chirp amplitude (constant or not).

By applying the Moyal formula to both terms of Eq. (15), we can write the statistic as the correlation of the combined WV $w_{1,2}(t,f) \equiv w_{y_1}(t,f) + w_{y_2}(t,f)$ of the synthetic streams with the template WV, viz.

$$\ell_{\tilde{p}}(\mathbf{x}) = \frac{1}{2N} \sum_{n=0}^{N-1} \sum_{m=0}^{2N-1} w_{1,2}(t_n, f_m) w_{s_{\tilde{p}}}(t_n, f_m). \tag{16}$$

This reformulation of the statistic turns out to be identical to in Eq. (9) or Eq. (14) if we replace $w_x(t,f)$ by $w_{1,2}(t,f)$. Therefore, the maximization of the time-frequency statistic goes along the same scheme exposed in the case of a single antenna.

When the source position θ and ϕ is not known in advance, we perform a full sky search by repeating the computation of $\max_{\tilde{p}\in\tilde{\mathcal{P}}} \ell_{\tilde{p}}$ for all sky patches of a celestial grid. If the maximum of the resulting "likelihood landscape" exceeds the detection threshold, we conclude that a GW chirp is present and the center of the corresponding sky patch yields the estimate of the source location.

Figure 7 present a proof of principle for the joint estimation of the source position and signal frequency. Here, the chirp amplitude is assumed to be constant.

We considered the simplified situation where the noise spectrum is flat. The realistic case where the noise spectrum is not flat different for each detector requires further approximations to obtain an additive statistic.[8] With those, the detection procedure globally holds.

4. CONCLUDING REMARKS

In this paper, we developed a methodology for the non-parametric detection of GW chirps. One of the main ingredients of the proposed method is the use of a finite family of template waveforms which "samples" the large and continuous set of admissible chirp signals. We construct the templates by chaining elementary bricks taken from a dictionary. The dictionary (chirplets) and the chaining rules (regularity constraints) are carefully chosen in order to capture the fundamental characteristics of the signal.

Following Ref. 5, we may say that we build a signal model from a graph: the template is associated to a path in the (chirplet) graph which describes the contents and structure of the admissible signal set. This approach does not apply to chirps only but it can be generalized whenever the signal (or the image) is composite in nature. For instance, we may consider paths (or subgraphs) in wavelet (or any other $\star$-lets) graphs (see Ref. 14 for an example with Markov trees of wavelet coefficients). We may trace the idea back to the Viterbi algorithm[15] used initially for error-correction scheme in digital communication. (In this case, the signal information is encoded a probabilistic Markov tree.)

The detection thus amounts to finding the best template/path with respect to a goodness of fit criterion. Dynamic programming is a well-known procedure to search for optimal path in a graph. However, it requires that the criterion to optimize be additive. Although this is not the case here, the good properties of the WV distribution allows us to extract the "additive part" of the exact criterion. The combining of dynamic programming and the WV distribution turns out to be a powerful time-frequency pattern search algorithm.

REFERENCES

1. E. E. Flanagan and S. A. Hughes, "The basics of gravitational wave theory," *New J. Phys.* **7**, p. 204, 2005. Available at http://arxiv.org/abs/gr-qc/0501041.
2. P. R. Saulson, *Fundamentals of interferometric gravitational wave detectors*, World Scientific, Singapore, 1994.
3. C. Cutler and K. Thorne, "An overview of gravitational-wave sources," in *Proceedings of GR16 (Durban, South Africa)*, 2002. Available at http://arxiv.org/abs/gr-qc/020490.
4. W. Anderson and R. Balasubramanian, "Time-frequency detection of gravitational waves," *Phys. Rev. D* **60**, p. 102001, 1999. Available at http://arxiv.org/abs/gr-qc/9905023.
5. E. Candès, P. Charlton, and H. Helgason, "Detecting highly oscillatory signals by chirplet path pursuit," *Appl. Comput. Harmon. Anal.* , 2007. In press. Available at http://arxiv.org/abs/gr-qc/0604017.
6. E. Chassande-Mottin and A. Pai, "Best chirplet chain: near-optimal detection of gravitational wave chirps," *Phys. Rev. D* **73**(4), pp. 042003 — 1–25, 2006. Available at http://arxiv.org/abs/gr-qc/0512137.
7. A. Pai, E. Chassande-Mottin, and O. Rabaste, "Best net-CC: Near-optimal coherent detection of unmodelled chirps with a network of detectors." In preparation.
8. O. Rabaste, E. Chassande-Mottin, and A. Pai, "Goniométrie pour les chirps gravitationnels," in *Actes du 21ème Colloque GRETSI*, (Troyes (France)), 2007.
9. H. L. Van Trees, *Detection, estimation and modulation theory – Part I*, Wiley, New-York, 1968.
10. E. Chassande-Mottin and A. Pai, "Discrete time and frequency Wigner-Ville distribution: Moyal's formula and aliasing," *IEEE Signal Proc. Lett.* **12**(7), pp. 508–511, 2005.
11. P. Flandrin, *Time-frequency/Time-scale Analysis*, Academic Press, 1999.
12. C. Papadimitriou and K. Steiglitz, *Combinatorial optimization: Algorithms and complexity*, Dover, New York, 1998.
13. H. L. Van Trees, *Optimum array processing*, Wiley, New York, US, 2002.
14. M. S. Crouse, R. D. Nowak, and R. G. Baraniuk, "Wavelet-based statistical signal processing using hidden Markov models," *IEEE Trans. on Signal Proc.* **46**(4), pp. 886–902, 1998.
15. G. D. Forney, "The Viterbi algorithm," *Proceedings of the IEEE* **61**(3), pp. 268–278, 1973.

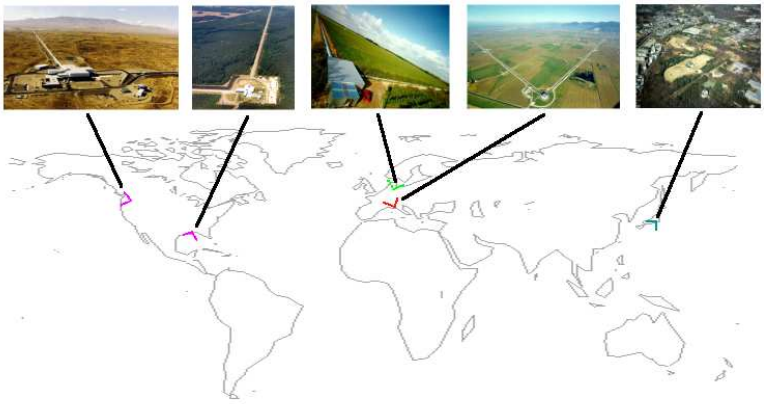

Figure 1. **The global network of laser interferometric GW antenna**. We show here the localisation and orientation of the antennas, along with an aerial view. The global network is currently composed of three antennas in the US (LIGO, `http://www.ligo.org`) [the northern one is a double antenna], two in Europe (GEO, `http://www.geo600.uni-hannover.de` [in green on the world map] and Virgo, `http://www.virgo.infn.it` [in red]) and one in Japan (TAMA, `http://tamago.mtk.nao.ac.jp`). GW antennas are Michelson interferometers whose arms (few hundred meters to kilometer scale) clearly form a L shape from above. In a nutshell, the measurement principle is the following. A laser beam (with a precise and stable wavelength) is divided into two twins, each of which being shown on one arm. The two beams are reflected by mirrors placed at the far end of the arms and then combines and interferes at the initial splitting point. The precise measurement of the resulting light intensity can be used to detect the tiny changes in the arm length, which evidences the contraction/stretching effect caused by an incoming GW.

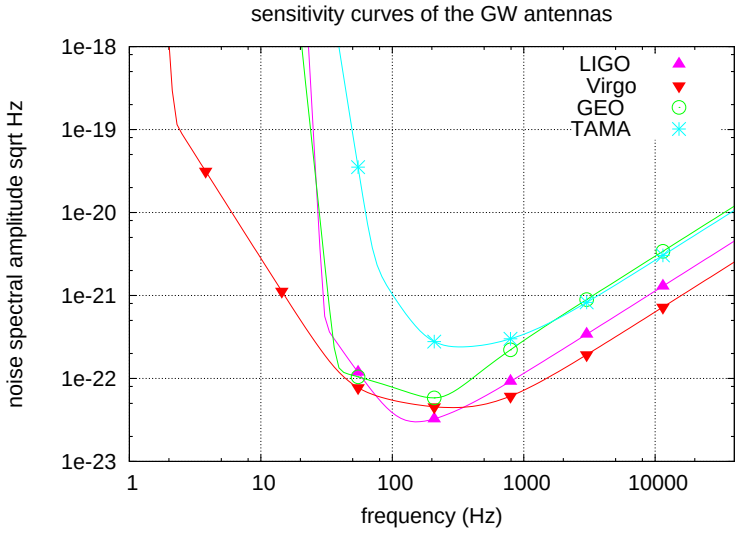

Figure 2. **Design sensitivity curves of the GW antennas**. This diagram presents the spectral density of the instrumental noise for the various GW antennas of the global network. This is also referred to as sensitivity curves (the lower the noise, the more sensitive the GW antenna). We show here the design curves and not the current status of the experimental ones. Several detectors are already running at design sensitivity.

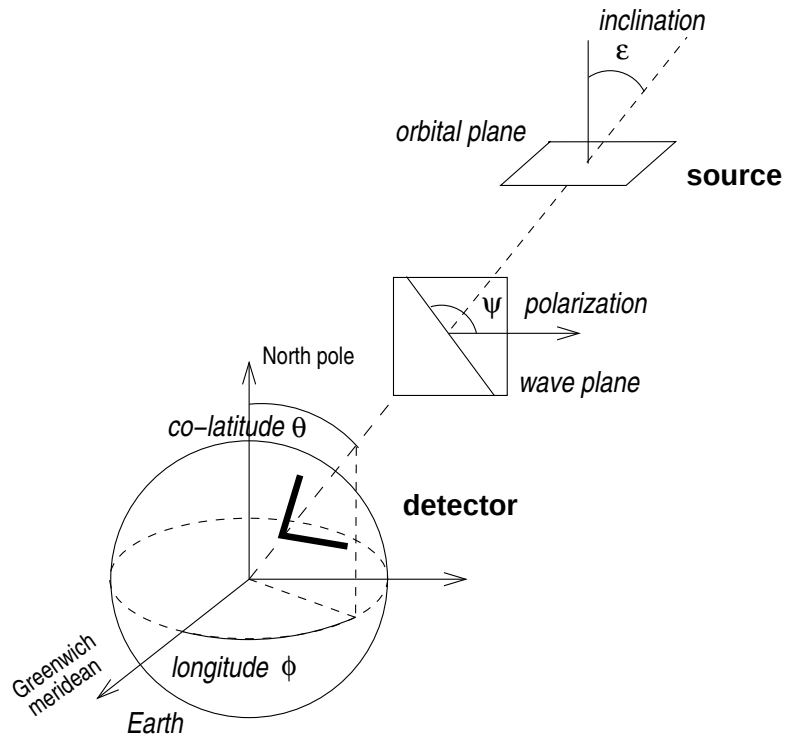

Figure 3. **Relative orientation of the source and GW detector and associated angles**. In the representation given here, the GW impinges the antenna with a normal incidence. In this (rather unlikely) situation, the antenna gain $|F|$ is maximized. The alignment is optimal for face-on sources (with an inclination angle $\epsilon = 0$).

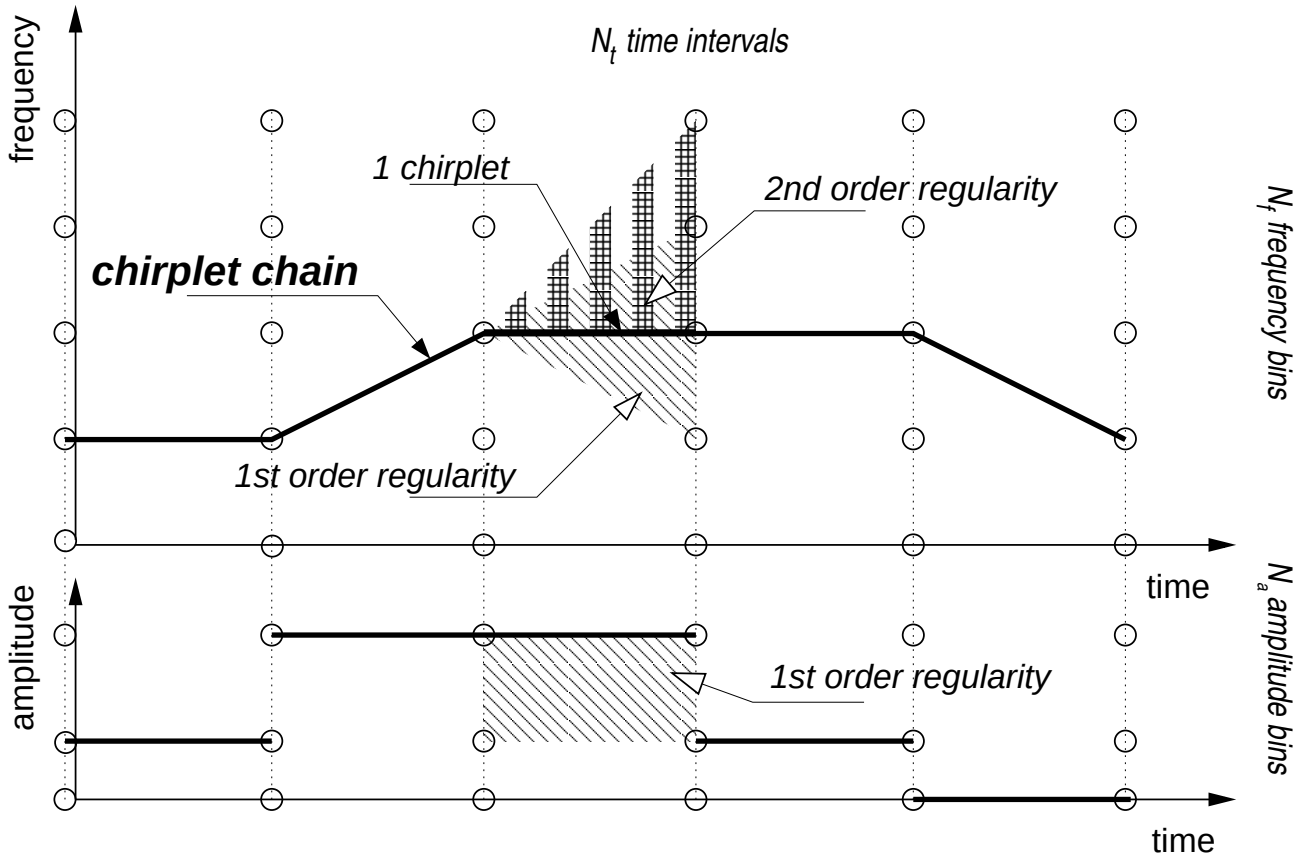

Figure 4. **Chirplets and chains of chirplets**. *Upper plot*: We refer to short linear chirps joining two successive nodes of the TF grid as *chirplets*. A chain of chirplets is a piecewise linear chirp and appears as a broken line. We limit the chirplet slope (domain associated to the 1^{st} order regularity) and the difference of successive slopes (domain associated to the 2^{nd} order regularity). *Bottom plot*: The chirplet amplitude takes values in dyadic scale (which includes also zero). The amplitude of the chirplet chain is thus piecewise constant. We limit the variations of the amplitude (as indicated by 1^{st} order regularity).

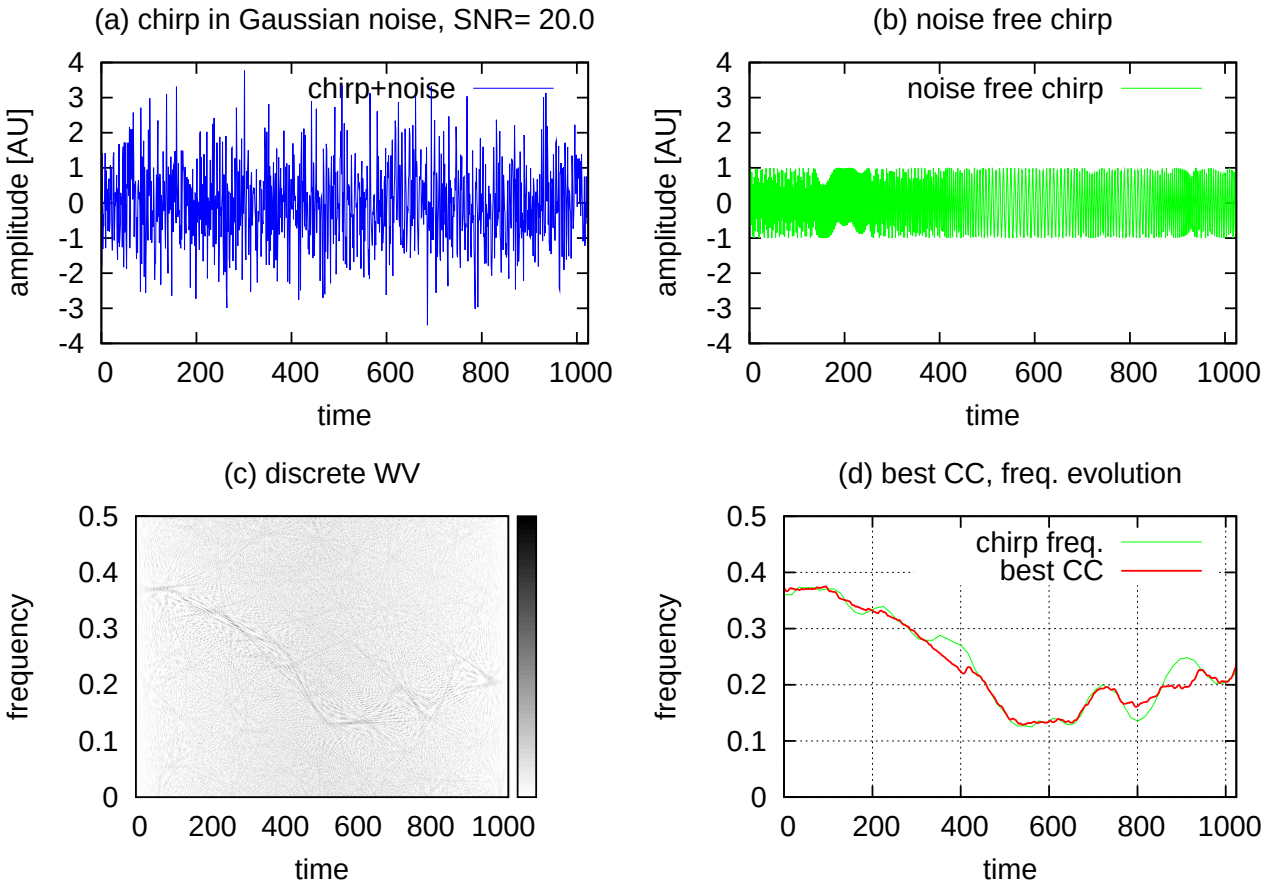

Figure 5. **Detection/estimation of a "random" chirp**. This is an illustration of the detection method described in Sec. 2.3. The test signal (b) is a "random" chirp (whose frequency follows a random walk in the time-frequency plane). We show in (a) and (c) the noisy chirp and its WV distribution. The signal to noise ratio is set to $\rho = 20$ with $\rho^2 = \sum_{k=0}^{N-1} s^2(t_k)$ (assuming a noise variance unity). In (d), we compare the exact chirp frequency $f(t)$ (thin/green) with the estimation $\tilde{f}(t)$ (bold/red) given by the maximum of $\hat{\ell}_{\tilde{p}}(x)$ and found by the proposed algorithm (best TF path in WV found by dynamic programming).

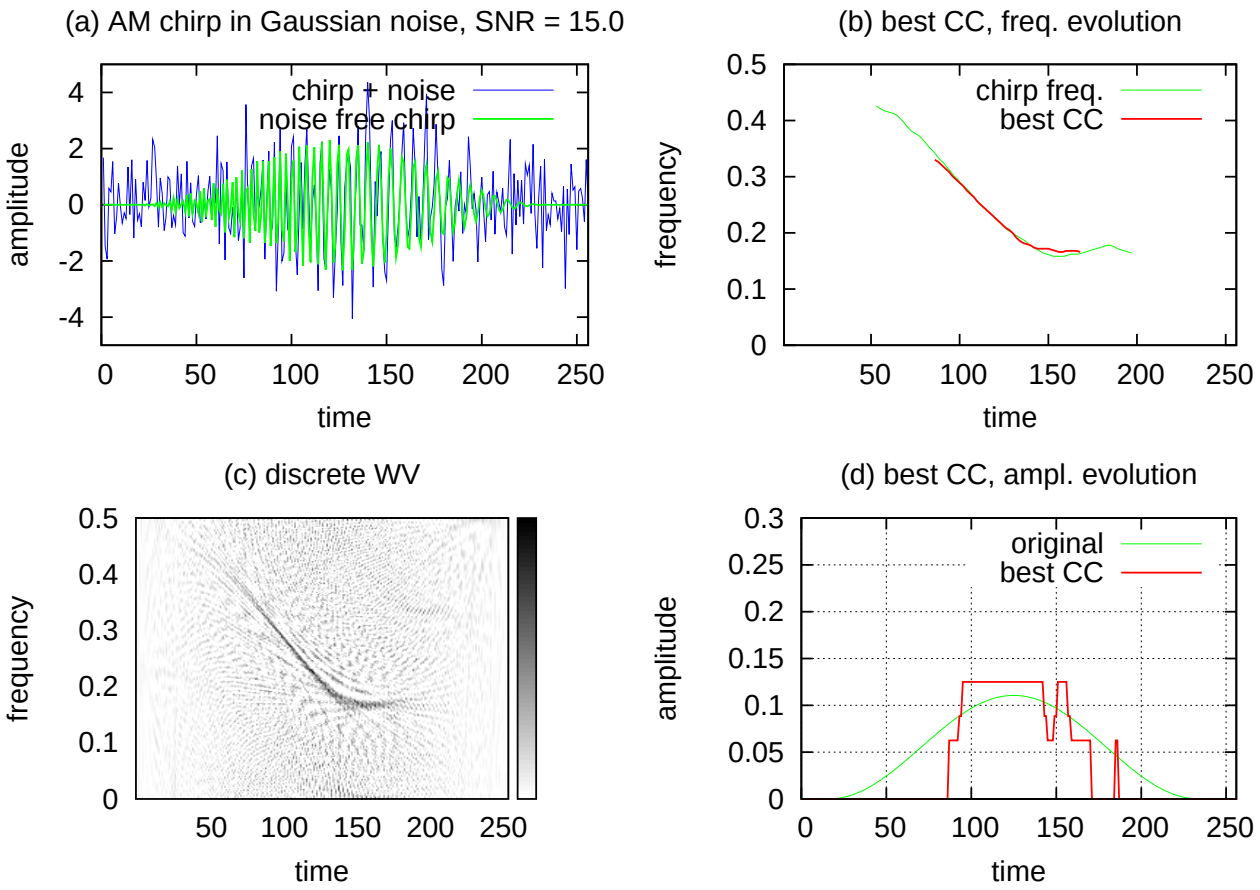

Figure 6. **Detection/estimation of a "random" chirp with amplitude modulation**. This is an illustration of the detection method described in Sec. 3.1. The test signal (a) is a "random" chirp with Hanning window as an envelope. (a) and (c) are the noisy chirp and its WV distribution respectively. The signal to noise ratio is set to $\rho = 12$. In (b) and (d), we compare the exact amplitude and frequency $A(t)$ (normalized to unit norm) and $f(t)$ of the chirp (thin/green) with the estimation $\tilde{a}(t)$ and $\tilde{f}(t)$ (shown only when $\tilde{a}(t) > 0$) given by the maximum of $\hat{\ell}_{\tilde{p}}(x)$ and found by the proposed algorithm. The time support of estimated amplitude appears narrower as compared to the exact one. This is partly due to the coarse sampling of the chirplet amplitude.

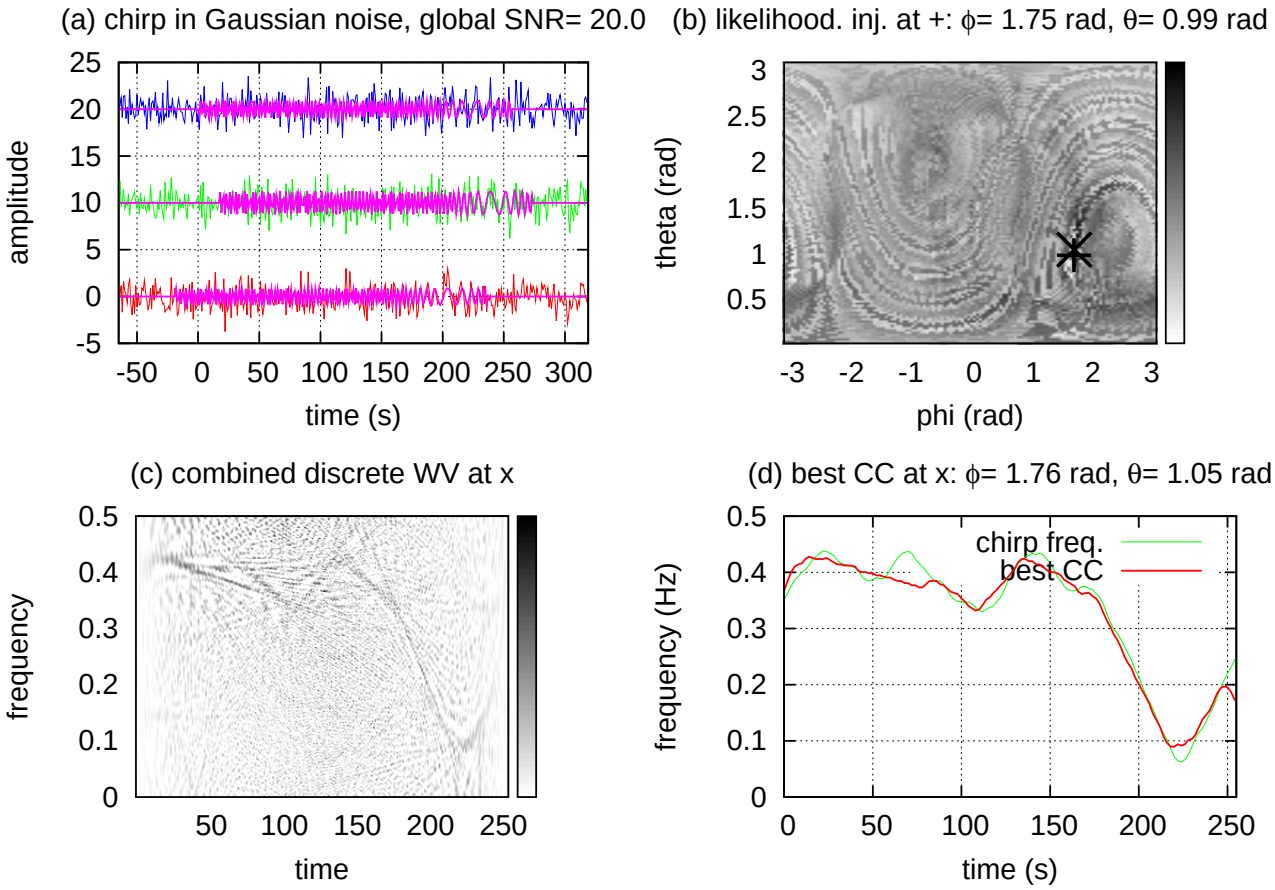

Figure 7. **Coherent detection/estimation of a "random" chirp with network of three GW antennas**. In the data of three GW antennas (the two LIGO and Virgo, see Fig. 1), we inject (a) a "random" GW chirp emitted from a source at the position marked with a "+" at $\theta = 0.99$ rad and $\phi = 1.75$ rad. The full sky search described in Sec. 3.2 produces a likelihood landscape (b) where we select the maximum. This is the detection point and it is indicated with "×". In (c) we show the combined WV distribution of the synthetic streams at the detection point. In (d), we compare the exact frequency of the chirp (thin/green) with the estimation obtained at the detection point.